Handling qualitative research outcomes

Papageorgiou, Spyridon N

DOI: https://doi.org/10.1080/14653125.2017.1392067

Originally published at:
Papageorgiou, Spyridon N (2017). Handling qualitative research outcomes. Journal of Orthodontics, 44(4):307-309.
DOI: https://doi.org/10.1080/14653125.2017.1392067

STATISTICAL CORNER

Handling qualitative research outcomes

Spyridon N. Papageorgiou

Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Plattenstrasse 11, CH-8032, Zurich

Correspondence to: Spyridon N. Papageorgiou, Clinic of Orthodontics and Pediatric Dentistry, Center of Dental Medicine, University of Zurich, Plattenstrasse 11, CH-8032, Zurich; T: +41 44 634 32 87 / F: +41 44 634 43 35; e-mail: snpapage@gmail.com

ORCID: 0000-0003-1968-3326

Words: 919

Theoretical scenario

Secure bonding of orthodontic appliances during fixed appliance treatment is closely related to treatment efficiency, as bond failures lead to increased costs in terms of time and money. Additionally, bonding and de-bonding of appliances at the respective beginning and end of the active treatment phase, and the subsequent removal of adhesive remnants from the teeth is both protocol-sensitive and one of the most time-consuming appointments in treatment.

Therefore, a group of orthodontists conducted a large-scale multi-centre clinical trial to compare the clinical performance of two adhesives, adhesive A and B (AA and AB, respectively), in terms of bond failure rate and type of bond failures. For now, we will focus mainly on the bond failure type. Bond failure type was evaluated for each failed bracket in a qualitative from standardized photographs using the Adhesive Remnant Index (ARI) (Kinch et al. 1989). ARI is based on an evaluation of the adhesive that remains of the tooth after the bond failure and uses a point-based system ranking of four categories, 0 to 3 (Table 1).

Table 1 around here.

At trial's end, the authors observed a total of 160 bond failures and assigned to each the corresponding ARI value after assessment of standardised photographs by two blinded orthodontists. The authors analysed descriptively the data by calculating the mean ARI in each group and its Standard Deviation (SD) and formally tested for differences between AA and AB with a Student's test for independent sample, considering a two-sided $P < 0.05$ as statistically significant. Their results are seen in Table 2.

Table 2 around here

The authors concluded that the ARI for both adhesives was similarly more close to 0 than to 1, which was confirmed by a non-significant P value from the Student's test ($P>0.05$). Therefore, no difference in the clinical performance of the two adhesives could be found.

Which of the following statements are correct, if any?

(A) The authors' correctly identified the nature of their outcome variable as a continuous one.

(B) Assuming the outcome was indeed a continuous variable, the author's analysis plan was correct.

(C) The author's transformed a qualitative variable to a quantitative variable.

(D) This transformation of the outcome variable couldn't have influenced the authors' conclusions.

The authors considered the outcome of their study (ARI) as a continuous variable. Continuous (quantitative) variables take numerical values and are measured on a scale, like weight in kilograms or height in centimeters. On the other hand, categorical (qualitative) variables have discrete categories, like male / female or like different colours. Categorical variables can be either non-ordered (like patient sex or country of residence), but they can also be ordered, which means that there is a progression in magnitude in the categories. An example of an ordered categorical variable is the education level of a group of people that can be categorised as no formal education, elementary school graduate, high school graduate, and college graduate. These categories can be ordered from lowest to highest in a logical order. The same can be done logically with the ARI categories and therefore, ARI values are indeed a categorical ordered variable, not a continuous one, and statement (A) is wrong.

However, assuming the trial's outcome was a continuous variable there might still exist potential issues to the authors' analysis plan. The statistical analysis plan can be affected by the nature of the dependent variable—i.e. if a variable is continuous or categorical. Additionally, the analysis plan is heavily dependent on the distribution of a variable—i.e. in our case if a continuous variable is normally distributed or not. There are certain ways to ascertain this both visually and formally through statistical tests, but the authors didn't employ any of them. If distribution diagnostics had been performed, then ARI – assuming this is a continuous variable – would present a highly asymmetrical distribution that is not compatible with normality. Therefore, the appropriate descriptive statistics would include the median and the interquartile range (instead of the mean and the IQR). Likewise, the appropriate inferential statistic to check for group differences would be the Mann-Whitney test (instead of the Student's test for independent observations). So statement (B) is also wrong.

Finally, the authors did indeed transform the qualitative (categorical ordered) variable of ARI in a quantitative (continuous) variable, as statement (C) correctly states. However, this has a profound influence on the trial's conclusions. If ARI had been correctly identified as a categorical ordered variable and had been analysed appropriately, then the appropriate descripitive statistics for it would include a cross-tabulation of the ARI categories' frequencies with the experimental groups in a so-called

contingency table (Table 3), also presented as a pie diagram (Figure 1). Formal inferential statistics then would include the Fisher's exact test, as is seen in Table 3.

Table 3 and Figure 1around here

We can clearly see that even though similar percentages can be seen for zero ARI between adhesives, adhesive B has significantly higher bond failures with ARI of 2 or 3, which means that a large portion of the adhesive remains on the tooth after failure. There may be several explanations for this difference in bond failure pattern. However, clinically speaking, assuming for now that AA and AB have a similar bond failure rate across patients, then AB might be less efficient for clinical practice, as a longer tooth clean-up appointment might be expected after de-bonding of the appliances. We can then conclude that the inappropriate handling of ARI lead to information loss and different conclusions—therefore statement (D) is wrong.

References

Kinch AP, Taylor H, Warltier R, Oliver RG, Newcombe RG. 1989. A clinical study of amount of adhesive remaining on enamel after debonding, comparing etch times of 15 and 60 seconds. Am J Orthod Dentofacial Orthop. 95:415–21.

Figure 1. Pie graph presenting the relative frequencies of each Adhesive Remnant Index (ARI) category amongst the bond failures of each adhesive.

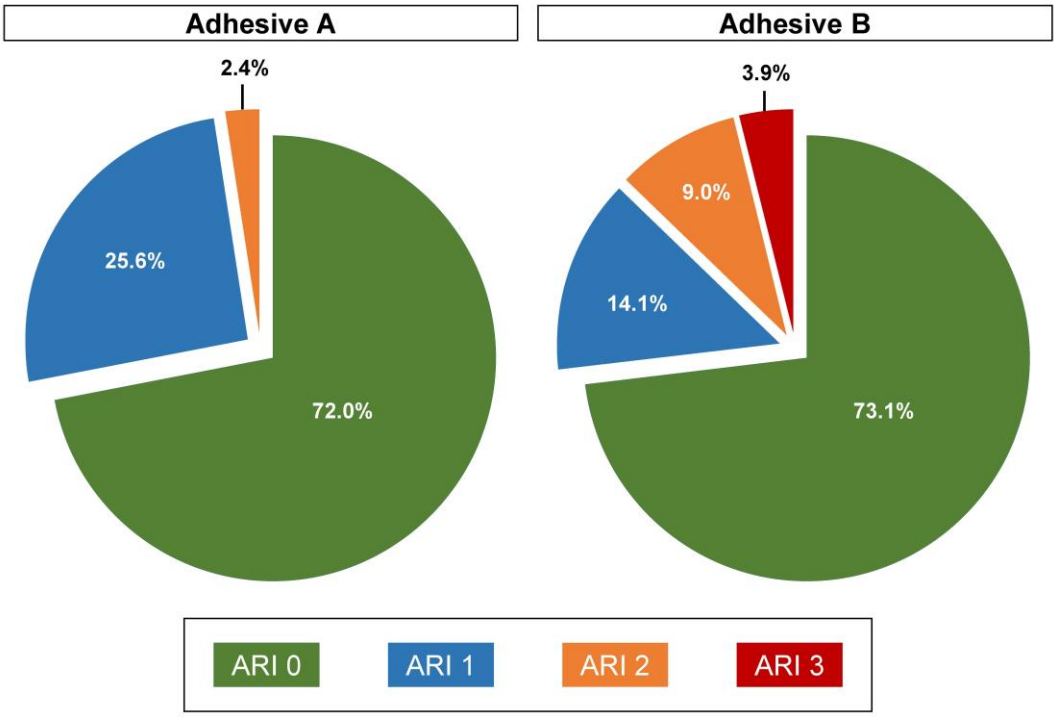

Table 1. Description of the ARI categories.

ARI category	Description
0	0% of adhesive left on tooth
1	<50% of adhesive left on tooth
2	>50% of adhesive left on tooth
3	100% of adhesive left on tooth; distinct impression of the bracket base

Adhesive Remnant Index, ARI.

Table 2. Results of the trial.

	Adhesive A	Adhesive B
Bond failure	82	78
ARI – mean (SD)	0.30 (0.51)	0.44 (0.82)
Student's test P value	0.22	

ARI, adhesive remnant index; SD, standard deviation.

Table 3. Contingency table with the results of the trial.

	Adhesive A	Adhesive B	Fisher's exact P
ARI 0	59 (72.0%)	57 (73.1%)	0.03
ARI 1	21 (25.6%)	11 (14.1%)	
ARI 2	2 (2.4%)	7 (9.0%)	
ARI 3	0 (0%)	3 (3.9%)	

ARI, adhesive remnant index.